REGION OF INTEREST CODING OF VOLUMETRIC MEDICAL IMAGES

Dimitris Agrafiotis, David R. Bull, Nishan Canagarajah

Image Communications Group, Centre for Communications Research, University of Bristol, UK

ABSTRACT

Three-dimensional wavelet coding of volumetric medical images provides better coding performance compared to corresponding 2D methods by exploiting the inter-slice correlation that exists in such data. It introduces however latencies when it comes to transmitting specific parts of the volume. This paper presents an extension to 3D-SPIHT which allows 3D Region of Interest (ROI) coding. ROI coding enables faster reconstruction of diagnostically useful regions in volumetric datasets by assigning higher priority to them in the bitstream. It also introduces the possibility for increased compression performance, by allowing certain parts of the volume to be coded in a lossy manner while others are coded losslessly. The necessary modifications to 3D-SPIHT for ROI coding are described and methods for specifying a 3D ROI without adding a significant overhead are suggested. Results are presented highlighting the benefits of the ROI extension.

1. INTRODUCTION

Digital medical image databases are very large with Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) accounting for a large proportion of the produced data. Effective storage and communication of such images requires coding which can adapt to the bottlenecks that arise in applications such as telemedicine and remote- Internet - browsing. This problem becomes even more acute for volumetric medical images due to the sheer volume of data. Moreover, storage of such data is problematic because of the requirement to preserve the best possible image quality which is usually interpreted as a need for medical images to be stored in a lossless manner. This requirement together with the need for fast transmission and browsing dictates the use of a coding method capable of progressive lossy to lossless coding. One such very well known method is the SPIHT algorithm [1]. SPIHT uses spatial orientation trees and successive approximation quantisation in order to code the coefficients resulting from a previously applied discrete wavelet transform. SPIHT produces an embedded code which enables the encoder or decoder to terminate the encoding/decoding at any point corresponding to a specific target bit rate. When reversible transforms are employed for the decomposition of the source image, like the S+P transform [2] or integer wavelet transforms [3], progressive lossy to lossless coding becomes possible, thus enabling the transmission of an image at various bit rates without the need for lossy compression of the original. This flexibility can be further enhanced by the introduction of 2D ROIs where greater "attention" is paid during the coding process [4] [5]. This can either mean that the ROI is coded losslessly while the rest of the image sustains some loss, or that the whole image is coded losslessly but the ROI gets higher priority in the bit stream. SPIHT works in 2D. For the case of volumetric medical images this implies that the underlying transformation is applied independently to each slice and hence the algorithm does not exploit the inter-slice correlation that exists in such images. Recently 3D extensions of SPIHT for lossless compression of volumetric medical images were introduced in [6][7] producing significantly better results than the original SPIHT. A 3D version of EZW [8] has also been presented [9] as well as 3D SPIHT for video coding [10]. Other 3D wavelet coding methods include those in [11] and [12]. In this work we present a ROI extension for 3D-SPIHT with the ROI being three-dimensional. First we briefly describe how to get a 3D version of SPIHT from the original. We then describe the necessary modifications to the 3D-SPIHT for ROI coding, based on an analogous 2D case [5] and suggest methods of specifying a 3D ROI without adding a significant overhead. Finally we demonstrate the coding benefits that the ROI extension brings to the transmission and storage of volumetric medical images.

2. FROM SPIHT TO 3D-SPIHT

The modifications needed to obtain a 3D SPIHT from the corresponding 2D version affect the applied transform, the spatial orientation trees and the adaptive arithmetic coding. A 3D instead of a 2D integer wavelet transform is applied to the volumetric image. As in the case of 2D separable wavelet transforms, 3D transforms are obtained by sequentially applying the 1D transform across all dimensions (one axial and two trans-axial). The decomposition applied to the source volumes unlike [6] and similar to [9] was fully dyadic. The spatial orientation trees used with 3D-SPIHT are 3D with each pixel (voxel) of a non-root non-leaf node having as offspring eight adjacent voxels in a 2x2x2 form one level below in the pyramid [10]. Arithmetic coding of the significance bits is also modified for the 3D case. Groups of 2x2x2 voxels (instead of 2x2 pixels) are coded with a single symbol from a number of different adaptive models conditioned to the number of insignificant voxels and descendant sets in the group. A total number of 164 models are used.

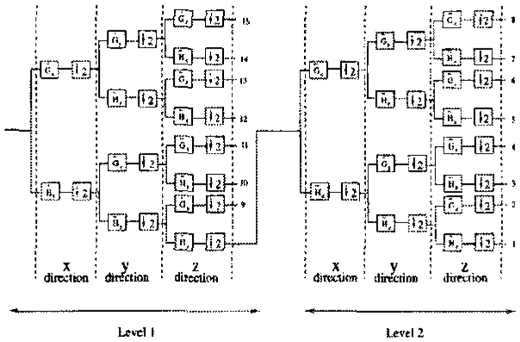


FIGURE-1: 3D dyadic decomposition with two levels.

One issue that comes up with the 3D version of SPIHT and the integer wavelet transforms mentioned in this work is the scaling that has to be applied to the transform coefficients in order for the transform to be approximately unitary. This is necessary in order to obtain a good rate distortion behaviour. As suggested in [2], the scaling factors that can be used are equal to $2/\sqrt{2}$ for the low pass coefficients and $1/\sqrt{2}$ for the high pass. For a 2D decomposition this results in multiplying each subband with factors of 2, which can be implemented as left bit shifts, and which preserve the reversibility of the transform. In the 3D case, however, the number of decompositions for each subband is not even unless a wavelet packet transform is employed as done in [6]. Scaling with the above factors for these subbands would result in coefficients which are not integer any more. However as suggested in [4],scaling by $2/\sqrt{2}$ and $1/\sqrt{2}$ the low and high frequency coefficients at each subband decomposition respectively is equivalent with scaling by 2 and 1, since multiplying both sets of coefficients with $\sqrt{2}$ doesn't change their relative importance.

3. ROI CODING WITH 3D-SPIHT

There are three issues that have to be addressed in order to add ROI coding functionality to SPIHT (both 2D and 3D). One is the selection of the wavelet transform coefficients that are necessary and sufficient for the lossless reconstruction of the ROI (the ROI lossless mask);the second issue is how to assign greater importance –priority- to these selected coefficients, and the third issue is how to specify the (3D in this case) ROI without adding a significant coding overhead.

3.1. Derivation of the ROI Lossless Mask

Although the region of interest is identified in the image domain it is the "wavelet image (volume)" – the coefficients of the wavelet decomposition - that is encoded and later decoded. Hence in order to assign a different priority to the ROI compared to the rest of the volume and/or guarantee the lossless reconstruction of the ROI, the coefficients that affect the ROI have to be identified [5][6]. To do that a three dimensional bitmap mask, named the ROI lossless mask, is maintained at the encoder indicating these coefficients. This mask undergoes a similar transformation to that of the forward wavelet transform of the image, where at each step the necessary coefficients for the perfect reconstruction of the ROI at the current level of decomposition are identified [5][7]. The inverse wavelet transform is consulted at each decomposition and based on the region of support of the synthesis filters (or the lifting steps of the inverse transform) those coefficients that will be necessary at the decoder for the perfect reconstruction of the ROI at the current level are selected. For example, assuming a 1-D one-level decomposition using the 5/3 integer transform, which has 3 low pass and 5 high pass synthesis taps, the low pass -L(n)- and high pass –H(n)- coefficients that are necessary for the reconstruction of the original samples X(2n) and X(2n+1) are L(n), H(n-1), H(n), L(n+1) and H(n+1).The 3D-ROI lossless mask is obtained with iterative application of the 1-D decomposition stages. It is obvious that the mask expands with each decomposition. In the 3D case this expansion takes place in the Z (slice) direction as well as the X and Y directions of the 2D case. Transforms with short filters (e.g. the 5/3) result in fewer coefficients being included in the ROI lossless mask which can lead to faster reconstruction of the ROI.

3.2. Specification of the ROI Priority

In order to assign greater priority to the ROI, the coefficients of the ROI lossless mask are scaled up through a fixed number (S) of left bit shifts (each left bit shift corresponds to scaling up by a factor of 2) [5]. The larger the number of left bit shifts, the greater the emphasis placed on the ROI and the faster the lossless reconstruction of it. As the algorithm checks for significant coefficients at each bit plane it will only code ROI coefficients for all the bit planes above the current (before the shifts) threshold. All trailing ROI zero bits at the last S bit planes are ignored by both encoder/decoder.

3.3. Description of a 3D ROI

The amount of additional information that needs to be sent to the decoder in order to enable decoding, is the description of the ROI and the speed of reconstruction (the number of shifts applied to the ROI). Having got this information the decoder can reconstruct the mask and scale the ROI coefficients down by ROI-shift bits. It is important for the coding performance of the algorithm that this additional information is kept compact. This means that the ROI is specified in a way which will not add a significant overhead to the final bit stream. When regular shapes are used for the description of the ROI the extra bits added to the coded file in order to describe the ROI are minimal. For example a cylindrical shape can be used to specify a circular ROI spreading to multiple slices. The centre, radius and height of the cylinder have to be included in the header of the coded volume adding minimal overhead to the bitstream. A flexible way of specifying a 3D ROI is through the use of volume cropping. A number of different three-dimensional ROIs can be created by means of specifying a minimum and maximum value for the X, Y and Z dimensions along with the type of cropping. Examples of ROIs created with each type of cropping can be seen in figure 3. The information that has to be included in the header is that of the min and max X,Y,Z values and the type of cropping used. For arbitrary-shaped ROIs (e.g. resulting from segmentation) one can use a method similar to the MAXSHIFT method specified in JPEG2000 [13]. By scaling the ROI coefficients so that the minimum coefficient magnitude belonging to the ROI is larger than the maximum non-ROI coefficient magnitude, the decoder can easily identify which coefficients belong to the ROI - any coefficient found significant during the first ROI shift bitplanes - and scale them down appropriately.

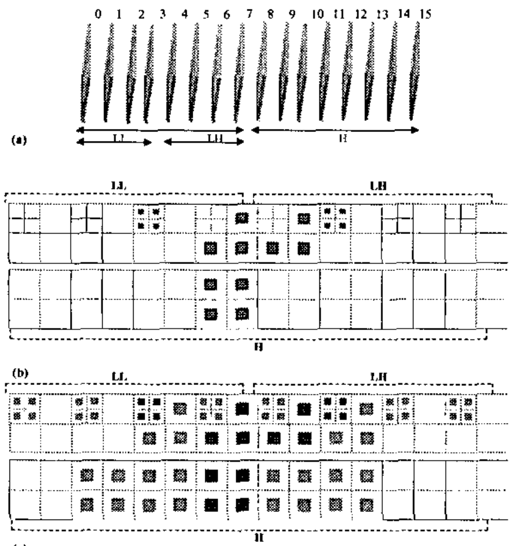


FIGURE-2 : ROI lossless mask for a square ROI of size 101x101 covering slices 6,7,8, of a 256x256x16 size volume using the integer 9/7M transform and a 2-level classic dyadic decomposition. (b) Spreading of the ROI in different subbands without expansion. (c) ROI lossless mask with extra coefficients shown in light grey.

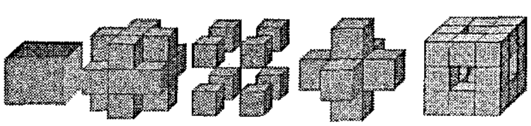


FIGURE-3: Different types of volume cropping for the specification of a 3D ROI. From left to right: sub-volume, fence, inverted fence, cross and inverted cross.

4. RESULTS AND CONCLUSION

Coding results are first presented for two volumetric medical datasets, one CT and one MR with sizes 256x256x128 and 256x256x64 respectively [14]. In both cases the 5/3 integer wavelet transform was used with 5 and 4 levels of decomposition. To assess the coding performance of the 3D-ROI-SPIHT with the CT dataset, a ROI was specified using the inverted fence cropping type with minimum and maximum values for the X, Y and Z dimensions of 145-255, 115-150, and 0-70 respectively. The created ROI includes all the voxels that do not belong in any of the above intervals and amounts to 21.6% of the volume. A 3D rendering of the ROI is shown in figure 4c. Results are presented with the ROI having a high priority (S=13) a medium priority (S=9) and a low priority (S=5), with 19 bits being the maximum significance threshold (including scaling for making the transform approximately unitary). Table 1 summarizes the results. The PSNR vs. bit rate performance for the ROI only and the whole volume are plotted in figures 4a and 4b. Similar results without ROI coding (S=0) are included in the graphs. As expected, 3D-ROI-SPIHT allows faster lossless reconstruction of the volume of interest at the expense of slower PSNR improvement for the rest of the volume. The inclusion of the ROI and the necessary shifting of the ROI coefficients result in a slight increase of the lossless bit rate for the whole volume. Similarly a ROI sub-volume of size 145x221x21 (16% of the volume) was specified for the MR dataset. The sub-volume can be seen in figure 5c. With the ROI having a moderate importance (S=8, maximum significance threshold = 17) lossless reconstruction of the volume of interest takes place at 0.52 bpp. A lower priority leads to slightly slower reconstruction. With volumetric datasets the imaged area is usually big including information which is not always diagnostically necessary. A typical example is that of MR images of the head where depending on the task ahead, the area of the brain might be the only diagnostically relevant area (the ROI). Assuming a prior segmentation of the volumetric image one could store the dataset at a rate that would ensure lossless reconstruction of the ROI and a high PSNR for the rest of the volume that provides the necessary anatomical context. An example of such a case is given below. 32 slices of a 256x256x160 sized MR volumetric image of the head are coded with the 5/3 integer transform and 3 levels of decomposition. The MAXSHIFT method is used to convey the ROI (the brain area) to the decoder, (S=16 with 15 being the maximum significance threshold). The ROI becomes lossless at 0.83 bpp with the volume PSNR being very low at this rate (15.97 dB). If the image is stored at 1.7 bpp then the average volume PSNR becomes 50.26 dB high enough to be considered visually lossless. The compression ratio in that case is 4.706 which compared to the standard 3D-

SPIHT (compression ratio 2.96, rate 2.69 bpp) represents a 63 % decrease in storage size. To conclude, we have demonstrated the benefits of adding ROI capability to 3D-SPIHT in terms of reconstruction speed of the volume of interest. We have also shown that if a small amount of loss can be tolerated for the background then 3D ROI SPIHT can greatly decrease the storage requirements for some volumetric datasets.

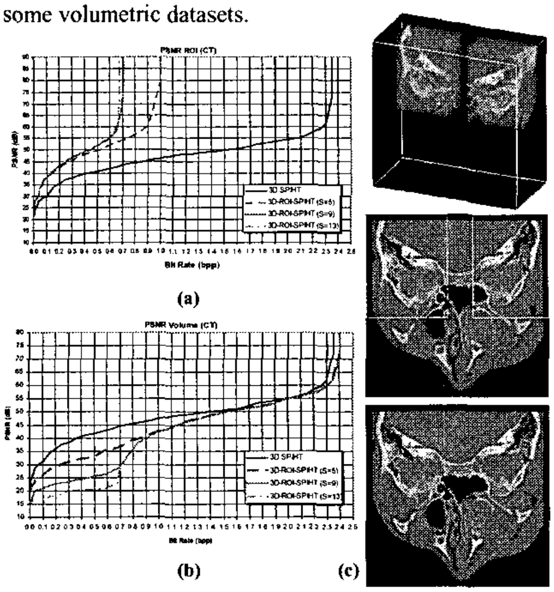


FIGURE-4: CT dataset: a) ROI PSNR and b) volume PSNR for the ROI shown at the top of the right column c) top - 3D ROI, middle - slice 78 of volume reconstructed at 0.72 bpp with lossless ROI marked, bottom - original slice.

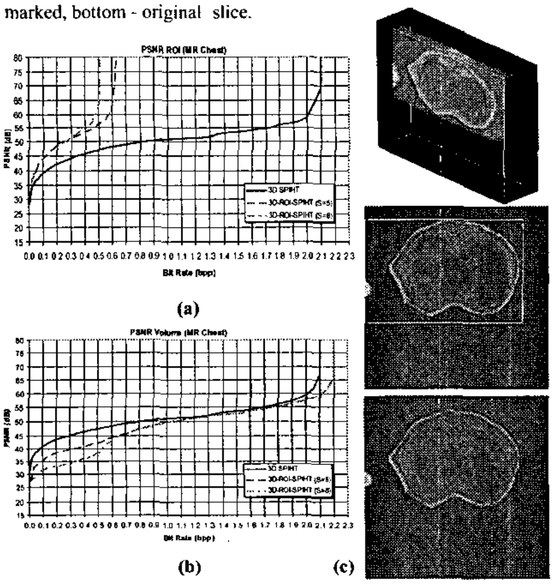
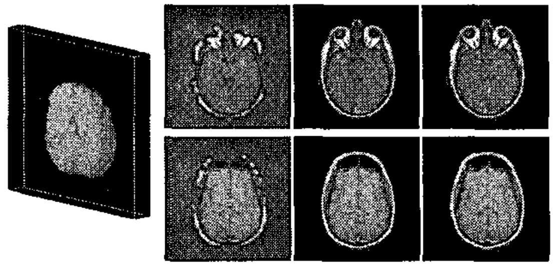


FIGURE-5 : MR dataset: a) ROI PSNR and b) volume PSNR for the ROI shown at the top of the right column c) top - 3D ROI, middle - slice 30 of volume reconstructed at 0.52 bpp with lossless ROI marked, bottom - original slice.



(a) (b) (c) (d)

FIGURE-6: Lossy / lossless compression of MR-Head dataset. a) 3D ROI b) 2 volume slices with ROI lossless at 0.83 bpp; c) same slices at storage rate of 1.70 bpp; d) lossless slices at 2.69 bpp

TABLE 1: CT Volume ROI results.

ROI Priority	ROI lossless rate (bpp)	PSNR at ROI lossless rate(dB)	Lossless bit rate (bpp)	Increase of bit rate (%)
S=0	2.360	Lossless	2.360	-
S=5	1.01	42.57	2.409	2.08
S=9	0.72	30.66	2.412	2.20
S=13	0.69	25.33	2.415	2.33

TABLE 2: MR-Chest Volume ROI results.

ROI Scaling	ROI lossless rate (bpp)	PSNR at ROI lossless rate dB)	Lossless bit rate (bpp)	Increase of bit rate (%)
S=0	2.124	Lossless	2.124	-
S=5	0.65	44.90	2.227	4.85
S=8	0.52	38.78	2.232	5.08

5. REFERENCES

[1] A. Said, W.A.Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees", IEEE Trans. CSVT, vol. 6, no. 3, 1996, pp. 243-250.

[2] A. Said, W.A.Pearlman, "Image multiresolution representation for lossless and lossy compression", IEEE Trans. on Image Processing, vol. 5, no. 9, 1996, pp. 1303-1310.

[3] A. R. Calderbank, I. Daubechies, W.Sweldens, and B.-L. Yeo, "Wavelet transforms that map integers to integers," Applied and Computational Harmonic Analysis, vol.5,no.3, 1998, pp. 332-369.

[4] D.Nister, C.Christopoulos, "Lossless region of interest coding", Signal Processing, no 78, 1999, pp. 1-17.

[5] E.Atsumi, N.Farvardin, "Lossy/lossless region of interest image coding based on set partitioning in hierarchical trees", ICIP 1998,87-91.

[6] Y.S. Kim and W.A. Pearlman, "Lossless volumetric medical image compression", Proc. of SPIE, vol. 3808,1999, pp.305-312.

[7] Z.Xiong, X.Wu, D.y.Yun, "Progressive coding of medical volumetric data using three-dimensional integer wavelet packet transform", IEEE Multimedia Signal Processing, 1998, pp. 553-558.

[8] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients",IEEE Trans.on Sig. Process.,vol.41,no.12,1993,3445-3462.

[9] A. Bilgin, G. Zweig, and M. W. Marcellin, "Three-Dimensional Image Compression using Integer Wavelet Transforms", Applied Optics, vol.39, No.11, 2000, pp. 1799-1814.

[10] B. Kim, Z. Xiong, W. Pearlman, "Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT)", IEEE Trans.CSVT, vol. 10, no.8, Dec 2000, pp. 1374-1387.

[11] G. Menegaz, L.Grewe, J.-Ph.Thiran, "Multirate coding of 3d medical data", ICIP 2000, Sept. 2000, pp. 656-659.

[12] P.Schelkens, A. Muntenau, J. Cornelis, "Wavelet coding of volumetric medical datasets", ICIP 2002, pp.225-228.

[13] C. Christopoulos, J. Askelöf, M. Larsson, "Efficient methods for encoding regions of interest in the upcoming JPEG2000 still image coding standard",IEEE Sig.Processing Letters,vol.7,no.9, 2000,247-249.

[14] Mallinckrodt Institute of Radiology, Image Processing Laboratory.